Article

new media & society
2021, Vol. 23(12) 3656–3676

DOI: 10.1177/1461444820960071
journals.sagepub.com/home/nms
SAGE

From archive cultures to ephemeral content, and back: Studying Instagram Stories with digital methods

Lucia Bainotti
Università degli Studi di Torino, Italy

Alessandro Caliandro
Università degli Studi di Pavia, Italy

Alessandro Gandini
Università degli Studi di Milano, Italy

Abstract

Despite growing interest, there is a shortage of research about the methods and challenges that concern researching ephemeral digital content. To fill this gap, the article discusses two research strategies to study Instagram Stories. These allow users to share moments of their everyday lives in a documentary and narrative style; their peculiar feature is ephemerality, as each Story lasts for 24 hours. The article (a) explores how to bypass the Instagram API closure and (b) engages in an attempt at 'circumventing the object of study', taking advantage of how individual users archive Instagram Stories on other platforms (here, YouTube). In so doing, we contribute to the debate that seeks to innovate and 'repurpose' digital methods in a post-API environment. Besides the methodological utility, we show the tension between ephemeral content and archive cultures, and raise epistemological and ethical concerns about the collection, analysis and archival of ephemeral content.

Keywords

Archive cultures, digital methods, ephemeral content, Instagram Stories, post-API

Corresponding author:
Lucia Bainotti, Università degli Studi di Torino, Lungo Dora Siena, 100 A, 10153 Torino, Italy.
Email: lucia.bainotti@unito.it

This article discusses two methodological strategies for the study of Instagram Stories. Released in August 2016, Instagram Stories is a feature of the social media Instagram that allows users to share 'all the moments of your day, not just the ones you want to keep on your profile' (Instagram, 2016). Stories are characterised by a documentary, narrative and everyday style, and enable users to combine the various modes of communication of the platform (pictures, videos, texts, emoji and stickers, audio) into a single digital object. Their peculiar feature is ephemerality, as each Story lasts for only 24 hours.

Despite the growing interest, there is a shortage of research about the methods and challenges that concern researching ephemeral digital content. Existing literature so far has largely provided a general understanding of ephemeral content, focusing on changes in temporality and documentation practices (Jurgenson, 2013), as well as on the uses of ephemerality in the context of self-presentation, intimacy and youth's sexual and relational cultures (Handyside and Ringrose, 2017; Koefed and Larson, 2016). In this context, the methodological strategies that are most often applied for the access, collection, analysis and archival of ephemeral content in the form of a 'story' rely on individuals' narrations and are collected through interviews and focus groups, often mixed with the observational analysis of online content. Yet, even when the focus is more centred on the content itself (Nashmi and Painter, 2018), current works largely avoid challenging the key issues of accessibility, epistemology and ethics that the study of this new means of communication entails.

To fill this gap, the article experiments and critically discusses two possible research strategies to study Instagram Stories as an object of social and cultural research. In doing so, we intend to contribute to the emergent literature that seeks to innovate and 'repurpose' digital methods (Rogers, 2013) in a post-API environment (Perriam et al., 2019; Venturini and Rogers, 2019). On the one hand, we explore a way to bypass the Instagram API closure, working around the restrictions to accessing Instagram data that have been enforced since April 2018.[1] On the other hand, we engage in an attempt at 'circumventing the object of study' (Bucher, 2017), taking advantage of how individual users archive Stories to ensure their permanence beyond ephemerality, and thus make them available on other platforms (in our case, YouTube). Besides the methodological utility, we show the emergent tension between ephemeral and archive digital cultures and discuss the epistemological and ethical questions that concern the modes of data collection, analysis and archival of ephemeral digital content. Our work is driven by two main questions: in which ways can digital methods be used for the analysis of Instagram Stories, and of ephemeral content more generally? Is the blending of qualitative and automated analysis a viable path for the study of Instagram Stories and ephemeral content?

The rise of ephemeral content: Instagram Stories

Despite the common understanding of the Internet as 'never forgetting' (Mayer-Schönberger, 2009), ephemerality has become a central component of many social media platforms, such as 4Chan (Hagen, 2018), Snapchat (Nashmi and Painter, 2018), and Instagram (Vázquez-Herrero et al., 2019). All of these platforms provide affordances designed to allow the creation of content which disappears from users' view after a short period of time. Existing research has looked at the diffusion of ephemeral content by

focusing mostly on Snapchat. Attention has been paid not only to the content of ephemeral snaps (Nashmi and Painter, 2018; Roesner et al., 2014), but also to how the temporality afforded by the platform may influence and mediate users' subjectivities and interactions (Bayer et al., 2016), with a particular focus on the implications for self-presentation (Koefed and Larson, 2016) and intimate relationships (Handyside and Ringrose, 2017). Moreover, these studies highlight that the ephemerality that the platform affords also allows for higher levels of users' self-disclosure (Bayer et al., 2016) and reduced self-presentational concerns, as compared to the more edited and permanent content posted on Instagram (Koefed and Larson, 2016).

Despite the increasing interest in ephemeral content, however, scarce attention has been paid so far to Instagram and its Stories feature, characterised not only by ephemerality but also by multimodality and everydayness. Unlike Snapchat, Instagram has represented for a long time the preferred platform to share static and persistent images, allowing users to share snippets of everyday life (Hu et al., 2014), selfies (Senft and Baym, 2015), or to perform micro-celebrity practices (Marwick, 2015), with important implications for brands and marketers (Carah and Shaul, 2016). Since 2016, following changes to the platform's architecture, the Stories format has become extremely popular, reaching the total amount of 500 million daily users in January 2019 (Newberry, 2019), and overtaking feeds as the primary way of sharing content (Constine, 2018). The rise of Instagram Stories reveals a tension between ephemeral content and archive cultures (Rogers, 2019) that is still overlooked in existing research. Handyside and Ringrose (2017) have already questioned the simplistic definition of Snapchat as a disappearing social media, by arguing that the platform offers 'an intriguing mixture of stickiness and transience, perceived permanence and elusive ephemerality' (p. 12). This is also the case for Instagram, as the in-built affordances to archive Stories allow for the possibility to save ephemeral content in ad hoc folders called 'highlights', which are visible on each users' profile. Furthermore, external strategies and tools for scraping and downloading Instagram Stories (e.g. StorySaver) have emerged together with cross-platform archival practices, particularly on YouTube. This can be a suggestion to consider the creation, duplication and storage of Instagram Stories as an evolution of the forms of representation and archival of *vernacular creativity* (Burgess and Green, 2018), as well as a by-product of Instagram ranking cultures (Rieder et al., 2018). In addition, it is interesting to question how ephemerality blends with another constitutive element of Instagram Stories, namely its storytelling dimension.

Exploring Instagram Stories: methodological strategies

From a methodological perspective, we build on the idea that Instagram Stories can be conceived as a kind of digital small stories (Page, 2015). These are a particular genre of online storytelling (conveyed through text, sound or image) where a poster can share 'mundane, ordinary and in some cases, trivial events' of their everyday life (Georgakopoulou, 2017: 268). Via digital small stories, Internet users have the opportunity to display before a digital audience a particular representation of themselves, a cultural point of view or a moral vision (Page, 2013). Unlike traditional, offline, small stories, digital ones configure as 'a-typical', that is, fragmented, open-ended and

intertextual. Social media represent privileged sites for the proliferation of small stories, insofar as their architectures constantly invite users to express themselves by means of micro-narrations (consider, for example, Facebook's prompt 'what's on your mind?'). Instagram Stories follow the same logic, only more explicitly, since they qualify as ad hoc devices for everyday (micro)storytelling that allow users to document mundane moments of their everyday life in real time.

Besides the fleeting ordinary moments that they allow us to see, Instagram Stories amount to be fleeting digital entities too. In fact, Instagram Stories last only a few seconds and, if not explicitly saved by the users on their own profile, they are no longer accessible after 24 hours. This ontological status makes these digital objects particularly challenging to study from a methodological perspective. That is why we hereby seek to devise ad hoc research strategies to locate, collect, analyse and archive ephemeral, story-format digital content such as Instagram Stories. The elusiveness of Instagram Stories not only lays in their content and format, but also in their technical status as data points. The Instagram APIs do not allow users to retrieve Instagram Stories. Thus, for example, the usually very useful Instagram Scraper (developed by the Digital Methods Initiative to collect Instagram posts by following hashtags and usernames, see Geboers, 2019) is not equally useful for exploring Stories. After the Cambridge Analytica scandal and the consequent curtailing of social media APIs (Bruns, 2019), Instagram Stories are among the social media data that cannot be (or can no longer be) obtained (Bruns, 2018). In the case of Instagram Stories, we can further speculate that this is due, on the one hand, to the private and personal information they contain or display, and on the other hand to the great business value they perceivably have (Puschmann, 2019), since a lot of interaction among influencers and regular users takes place via Instagram Stories (Warren, 2019).

In order to address the methodological challenges that Instagram Stories raise, our strategies turn to *follow the actors* (Latour, 2005), as we take advantage of the *natively digital methods* (Rogers, 2013) by which Internet users capture and archive Instagram Stories themselves (Postill and Pink, 2012). In so doing, we seek the users' collaboration, making them, *de facto*, our co-researchers (Caliandro and Gandini, 2017). Drawing on this epistemological principle, we hereby explore two methodological strategies, specifically consisting in: (a) bypassing the Instagram APIs and (b) circumventing the object of study.

Strategy 1: bypassing the Instagram APIs

Our first strategy consists in bypassing the Instagram APIs by using scraping techniques (Cooley et al., 1997). Scraping is an IT technique that enables researchers to grab specific digital entities (e.g. the title of an article) directly from the HTML code of the webpage in which they are located (Weltevrede, 2016). Scraping is a controversial topic in academic research, of which the legal and ethical contours are fuzzy (Landers et al., 2016). Although not illegal per se (Waterman, 2020), scraping is a practice to which social media platforms are particularly averse, for a number of reasons. First, in order to pull the whole content off a webpage, a script has to make a large number of 'calls' to the server hosting the webpage. If this process is repeated multiple times, by multiple users, on multiple pages, it carries the concrete risk of engulfing the server and, eventually,

crashing the website – a risk that social media companies want to ward off at all costs. Second, scraping permits to sneak into private profiles and access information that users are not necessarily willing to share. Last but not least, by using scraping techniques, developers and researchers may access data that social media companies do not intend to share, as their business model is mainly based on the selling of such data (Zuboff, 2019). Yet, as argued by Venturini and Rogers (2019: 536–537), forms of scraping are a 'necessary evil' for social research if performed conscientiously. In their view, scraping 'forces researchers to observe online dynamics through the same interfaces as the actors they study' (Venturini and Rogers, 2019), and thus take advantage of the ways in which users themselves generate or manage content.

In our case, we used a freely available tool to scrape Instagram Stories – that is, StorySaver.[2] StorySaver allows to visualise the Stories posted by users with public Instagram profiles, and to download them as .jpg or .mp4 files. The tool replicates the ephemerality of the object of study, as it collects the Stories within the 24-hours of their permanence on each user's profile. The Stories collected thanks to StorySaver were then organised in an anonymous corpus.

Ethical considerations about Strategy 1

So as to account for the controversies concerning scraping, we decided to employ for Strategy 1, one of the many free online tools for capturing public Instagram Stories. Therefore, we used a scraping technique without programming our own scraping script in order to (a) bypass the platform's restrictions or blocks (Chellapilla et al., 2005), (b) disguise the non-human identity of the collector of data (Von Ahn et al., 2003), or (c) access to content that is protected by privacy settings or passwords (Franzke et al., 2019). In this way, we are able to comply with the Terms of Service[3] of Instagram (Fiesler et al., 2016). Second, we treated data in ways that caused no harm to users and were respectful of their privacy. We collected data by following 'neutral' Stories (e.g. related to everyday errands) and avoiding those dealing with sensitive topics (e.g. political views, or sex). Moreover, from 2018, users can activate the function of 'close friends', which makes their Stories visible to only a list of chosen friends, hiding them instead from all other followers; this reduces the probability to run into content that users deem intimate. Given the difficulties often related to the request of consent in social media research, we consider the data collected as similar to that gathered from observation methods (Light et al., 2018). As a consequence, we will not display any screenshots regarding the Stories in this paper, nor share usernames, links to individual profiles or any other personal information. Moreover, we analysed data in an aggregated as well as clustered form, and presented them through coding categories. Such categories are very general (such as 'portrait', 'landscape', 'mood', etc.) and, consequently, do not allow for the identification of users. Finally, we did not share our dataset with third parties, since it was gathered exclusively for academic purposes. In short, we did not incur in the typical ethical issues that are normally ascribed to scraping techniques, since we did not '[break] the law, or [put] a burden on a site's servers, or potential(ly) harm [. . .] users' (Fiesler et al., 2020: 10).

Although the procedures described above comply with general ethical and legal protocols, another matter of concern remains. Users are not always aware that their public

data can be used for research purposes and might expect it to remain 'private' (Zimmer, 2010). Literature commonly concurs that this happens because, customarily, users do not read carefully the platforms' terms and conditions (Böhme and Köpsell, 2010) – and also because these are often extremely long and difficult to comprehend (Reidenberg et al., 2015). Nonetheless, the fact that regular users tend to disregard platforms' terms and conditions cannot lead us to take for granted that they ignore, automatically, that researchers might analyse web data. In fact, as Landers et al. (2016) have remarked, the recent privacy controversies (i.e. Cambridge Analytica) 'have increased awareness that any data shared over the Internet has been in effect shared publicly' (p. 487).

With reference to our specific case, we acknowledge that capturing Instagram Stories clashes with the conceptual notion of ephemeral content. It may be questioned whether researchers are actually legitimised in the first place to access, collect and archive for research purposes content that users produce in the understanding that it will soon become unavailable. It may be even pointed out that, compared to a traditional social media post, ephemeral content might be seen as a more privacy-friendly practice of posting by many users, as a way to publish material that they do not want to remain permanently available on the platform. We contend this argumentation does not hold, for two reasons. First, while many users might assume that the ephemeral content they are posting will be erased, this is not expressly guaranteed anywhere in the platforms' policies. Second, as we will see in the section on YouTube, users seem to deem somewhat culturally valuable to duplicate certain ephemeral content – and we as researchers followed the same principles. Therefore, despite its temporality, we treated ephemeral content just like any other content, upholding to the same ethical standards and privacy prescription. Finally, we consciously adopted a data-activist stance. As argued by some of the most prominent digital scholars (Bruns, 2018; Kazansky et al., 2019), in an era of APIs curtailing, social media researchers must engage in devising new and alternative methods to keep social media research alive; this, together with the ultimate goal of keeping critical thinking alive, which is something that this research wishes to promote (Bruns, 2019).

Data collection

Before starting our inquiry, we needed a point of entrance in the Instagram Stories ecosystem, so as to avoid collecting data in an overly random and subjective way. Thus, we accessed the Instagram platform by following the generic hashtag #happy. First, the choice of such a generic hashtag allows us to run into users' ordinary content and, second, it lowers the risk of incurring in sensitive topics that could have caused harm to users. After setting the keyword, we launched Instagram Scraper (DMI), which retrieved 10,000 posts containing the hashtag #happy (March 2019 to April 2019). Then, we calculated the users' distribution for the number of published posts. This operation allowed us to distinguish between the very active users and the less active ones. We concentrated our attention on the long tail and on that majority of users (n = 6.399; 88.80%) who had posted only once. We did so because, as mentioned earlier, we were interested in studying Stories that had been posted by ordinary users – and, thus, avoid influencers and bots. In line with our operationalisation of the notion of 'ordinary user', we randomly extracted from this pool a number of profiles following two main requisites: the profiles (a) are not

bots, fan pages or brands, and (b) have less than 2000 followers. Our final sample consisted of the first 15 randomly extracted users who matched these characteristics. Then, each user in the sample was followed for 7 days (25 April 2019 to 1 May 2019) and their Stories were collected every day of the week at the same time, in order to account for their daily Stories-sharing activities. Through this procedure we created a corpus of 292 Stories. The resulting empirical materials were then organised in a spreadsheet containing an arbitrary id label for each Story, the date of extraction, an anonymised user id, together with the transcription of the texts and audio content of each Story.

Data analysis

The scraping procedure through StorySaver only allows collecting Stories from a single user. This peculiarity, together with the ethnographic and manual work that is required to gain the empirical material, does not allow for the creation of an extended dataset. Given these limitations, we analysed the visual and audio content of the Instagram Stories by means of an ethnographic coding approach (Altheide, 1987), which blends non-intrusive participant observation and note-taking with coding practices from the content and visual analysis traditions. More specifically, the analysis builds on the integration of compositional and content analysis (Rose, 2016), paying specific attention to both still and moving images (Heath et al., 2010). Given the exploratory nature of the study, and in line with the principles of an ethnographic coding approach (Altheide, 1987), the existing literature on Instagram (e.g. Leaver et al., 2020) initially guided the creation of the visual codes, while other descriptive and analytical labels were expected and allowed to emerge throughout the study.

Moving from the *denotative* to the *connotative* level (Banks, 2007), the analysis aims at grasping the visual content and practices of use of Instagram Stories, and consists in five steps, summarised as follows:

Denotative level:

- Format
- Visual codes

Connotative level:

- Narrative style
- Context of use
- Grammars

We started the analysis by focusing on the denotative level and visual content of the collected Stories, taking each Story as a single unit of analysis (Rose, 2016). In order to account for both still and moving images, we started analysing the ‘format’ of each Story, distinguishing between static pictures (66%) and videos and small animations (34%). Given the short duration of Instagram Stories (max 15 seconds), we noticed that the videos we collected could be considered as the transposition of a singular event in a dynamic format. For this reason, we followed a similar procedure for the visual content analysis of both static and dynamic content.

Second, we delved into the content of each Story accounting for the predominantly visual codes. In particular, we looked at two types of codes: (a) *content-related codes*, addressing what is represented in each photo/video and (b) *Instagram-specific digital objects*, that is, the visual elements inscribed within the platform affordances that allow to create and beautify Instagram Stories – specifically stickers (e.g. emoji, gif) and what we call 'interactive stickers' (e.g. poll; slide emoji stickers[4]). We then analysed and coded all the Stories in the dataset, describing the predominant visual elements while looking for patterns, similarities and differences within the corpus. After open-coding each video and photo, we grouped the denotative labels into broader visual content categories (Figure 1).

As Figure 1 shows, the most recurring visual components represented in our corpus refer to the category 'portrait' (33.90%) and mostly show users themselves (labelled as 'selfie', 20%), or alternatively friends (7%) and family (3%). Instagram Stories also display material assemblages (indicated with the category 'materiality', 12.35%), characterised by the presence of both material objects (6%) and body parts (4%). Some other recurring visual elements are the representations of specific settings and landscapes (10.95%), celebrations (9.93%) and food (9.24%). Both of the categories 'portraits' and 'materiality', two of the most recurring ones, reflect the most common visual elements represented in Instagram posts (Hu et al., 2014), pointing to the fact that some visual codes that are considered as Insta-worthy are shared persistently, despite their ephemerality. The ways in which these objects are photographed and combined recalls the presence of specific Instagram aesthetics, such as the many ways of representing brands as assemblages (Rokka and Canniford, 2016) and other conventions in displaying selfies (Hess, 2015).

Besides these predominant visual elements, the second most recurring visual category goes under the name of 'composition' (22%). The Stories within this category are characterised by the juxtaposition of different elements, mostly text (10%), stock images (5%) and memes (2%). Notably, in this type of Stories, the main focus is not a visual element, rather a composition of texts and images created by the user. In this case, it is the juxtaposition of different elements that guides the focus and meaning of the Story. Moreover, looking at the distribution of digital objects, one can notice that, despite the ubiquitous presence of GIFs and emojis in almost every Story, 'interactive stickers' become an essential element in the composition category (3%). Digital elements inscribed within the platform's affordances are thus combined with more traditional visual elements, creating a composition that is peculiar to the Stories-format. This multimodality is one of the main specificities of Instagram Stories, which clashed with the 'classic', more static, Instagram posts.

In the second part of the analysis, we moved to the connotative level and examined the Instagram Stories in relation to their broader cultural meanings. First, we considered the 'narrative style' of the Stories in our corpus. In line with the definition of 'small stories' and their peculiarities previously outlined, the results show that the majority of the empirical material consists of Stories as single units (60%), namely, single snippets with their own meaning. The remaining 40% is organised in coherent narrations, documenting a given event by means of different frames that are organised in a slideshow, rather than articulating a linear story. Despite the presence of some micro-narrations in a very documentary style, each Story maintains a meaning of its own, and its content as well as its

Categories	Occurrences (%)
Portrait	**33,90%**
Selfie	20,06%
Friend	7,49%
Family	2,99%
Composition	**21,57%**
Text	10,18%
Images	5,09%
Interactive sticker	2,99%
Memes	2,40%
Drawings	1,20%
Materiality	**12,35%**
Material objects	5,99%
Bodies	3,59%
Car	2,10%
Books	1,20%
Setting	**10,95%**
City	6,29%
Nature	5,39%
Travel	2,40%
Celebration	**9,93%**
Party	6,59%
Special occasion	3,89%
Food & Beverage	**9,24%**
Food	4,79%
Drink	3,29%
Others	**1,02%**
Animals	1,20%
Various	0,90%
Total	**100,00%**

Figure 1. Visual content analysis – coding categories.

contextual use can easily be understood without taking into consideration the whole narration in which it is set.

Therefore, moving to the interpretation of Instagram Stories' 'context of use' we analysed each video and image in itself, while taking into account the other Stories shared by the same user as contextual elements. This procedure resulted in the creation of 7 categories representing different ways of interpreting Instagram Stories as a social practice (Figure 2).

Context of use	Occurrences (%)
Special event	37,33%
Daily life	25,68%
Mood	10,62%
Interaction	9,59%
Self-display	7,88%
Aphorism	5,48%
Other	3,42%
Total	**100,00%**

Figure 2. Instagram Stories – context of use.

Unsurprisingly, the vast majority of Stories accounts for what we labelled as 'special events' (37.33%), namely those moments considered by users as out of the ordinary, and thus worthy of being recorded and shared. The second most popular category includes snaps representing moments of everyday life (25.68%). Moreover, the aim of a certain number of Stories is to share an image of oneself ('self-display', 7.88%). Two other categories are worthy of attention, what we called 'mood' (10.62% – stories expressing personal feelings) and 'interaction' (9.59% – stories that attempt to establish a flow of communication with followers).

Finally, by matching the visual codes of the analysis (Figure 1) with the context of use (Figure 2), we were able to identify some specific 'grammars' (Figure 3). The concept of 'grammar' points to a set of aesthetic norms that characterise Instagram Stories, as well as to the possible actions that the platform offers to users (Gerlitz and Rieder, 2018). In this sense, the idea of grammar takes into consideration the visual elements and how they are combined in their compositional modality (Rose, 2016), the cultural meaning they reflect, and the role of the platforms in prompting them.

Figure 3 shows the different types of visual elements associated with specific Instagram Stories' context of use (just the four most recurring ones are represented). It emerges that the narration of special events is mostly represented by means of portraits (44.95%) and celebrative moments (25.96%), whereas the 'daily life' context of use is characterised by the presence of material assemblages (32%) and the representation of everyday settings (24%). Yet, we can see that users communicate their mood and try to create interaction mostly by means of the 'composition' type of Stories (74.1% and 60.71% respectively), thus relying on the juxtaposition of different visual and textual elements. According to these results, two main grammars can be outlined: a *grammar for documentation* and a *grammar for interaction*. The first one is related to the celebration of special events as well as to the representation of ordinary, everyday moments, and it is based on portraits and materiality. These results support the idea that Instagram Stories promote the visual display of ordinary life in a context characterised by the ubiquity of photography. With their Stories, regular users seem to reaffirm the original purpose of Instagram as a social media for sharing pictures taken on-the-go, which has recently been challenged by the presence of staged and polished content, mostly posted by celebrities and influencers.

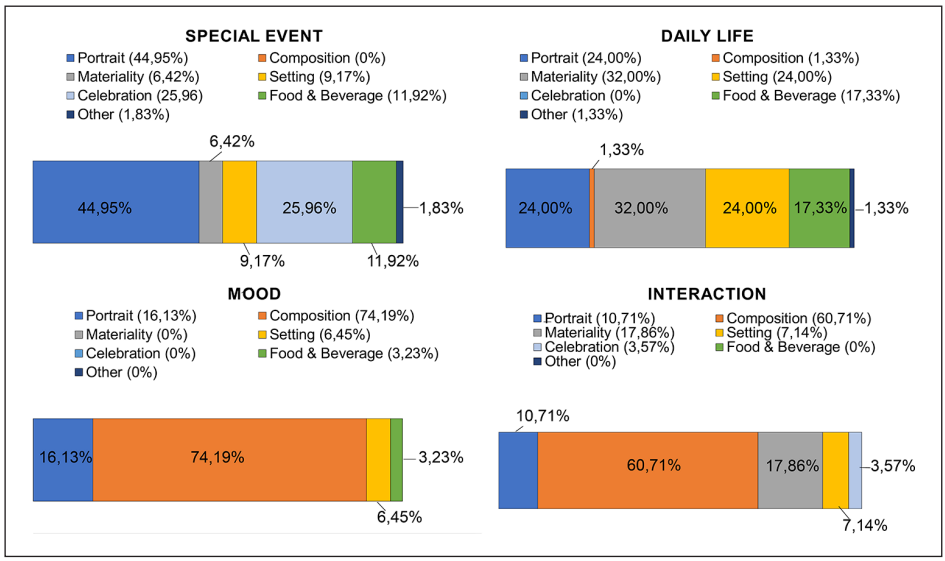

Figure 3. Visual categories distributed per context of use.

The second grammar, labelled as 'grammar for interaction', is mostly characterised by the 'composition' type of Stories, which are used to share one's mood and to generate interaction. This grammar seems particularly influenced by the indications of how-to tutorials that suggest the best ways to create engaging content, visibility and revenues (more on this later). In particular, with the 'compositional' type of Stories, users tend to mimic the practices adopted by celebrities and influencers in a well-known form of micro-celebrity transposed to the Stories-format (Marwick, 2015). Despite the adoption of codes and practices related to marketing content, however, these attention-seeking practices seem less directly aimed to outsmart the Instagram algorithm but, rather, to recall forms of what Bucher (2018) calls 'programmed sociality'. Such a concept is useful to explain how communication flows and how interactions are oriented by the platform's affordances through Instagram Stories and at the same time measured and evaluated according to the ubiquitous criterion of engagement. In this sense, the grammar for interaction can be considered as one of the means by which the platform's affordances function to encourage users to connect and engage with each other.

This kind of analysis sheds light on some of the visual peculiarities and the contextual use that characterise Instagram Stories. Because visual discourses significantly change according to the hashtag under analysis, these results might be too specific to be generally applied. Still, our study is able to demonstrate that the Story multimodality can assume a documentary style, and create flows of communication and interaction among users. These elements, together with the ephemerality of the medium, represent the peculiarities of Instagram Stories as compared to the traditional Instagram feed. Focusing instead on the aesthetics of the visual texts in our corpus, despite the presence of Stories-specific grammars, some compositional norms related to the broader platform vernacular (Gibbs et al.,

2015) still persist. Not only are traditional Instagram-worthy elements represented, such as selfies and assemblages of material objects, the ways in which they are depicted and framed are also in line with already-established Instagram aesthetics (Manovich, 2016). Thus, despite the feature of spontaneity embedded in the Stories' ephemerality, the moments captured continue to reveal some platform-specific compositional norms. The ephemerality of Instagram Stories thus allows users to share more spontaneous content, which is nonetheless somehow coherent with broader Instagram aesthetics.

Strategy 2: circumventing the Instagram platform

An alternative strategy for the study of Instagram Stories is the possibility to 'circumvent the object of study'. Due to the peculiar nature of ephemeral content, which is destined to disappear, we assumed that this was prone to practices of duplication, storage and archival by users who would want to preserve it because they consider it somewhat culturally valuable or to simply keep it available beyond its 'expiry date'. Accordingly, we expected to find Instagram Stories replicated and stored outside the Instagram platforms, particularly where audio-visual 'vernacular' content is already commonly archived and made available to others for sharing and commenting. For these reasons, we turned our attention to YouTube, that is, widely considered to be *the* digital repository of present-day popular culture (Burgess and Green, 2018).

The strategy of 'circumventing the object of study' is not new in digital social research. Here, we take inspiration from Gerrard (2018), who located pro-anorexia communities on Instagram, Pinterest and Tumblr by 'circumventing' hashtags (such as #proana or #thinspiration) as a mechanism of search (since members of those communities explicitly avoid using hashtags in their messages), and particularly from Bucher (2017), who used tweets by ordinary users who reflected on the Facebook algorithm in order to better understand how the latter works. While these tweets did not offer technically accurate information about the algorithm itself, they nevertheless allowed Bucher to grasp the social imaginary that ordinary users construct around it.

In our case, we circumvented the Instagram platform by searching for Instagram Stories on YouTube. This strategy aims at accessing Stories that users duplicate and store with the implicit assumption of their relevance. In this sense, here we *follow the actors* in a way that is practically different, but epistemologically analogous to Strategy 1. The use of YouTube to 'circumvent the object of study' and thus search for Instagram Stories outside the Instagram platform has several advantages. First, it allows us to observe Stories that do *not* expire after a set time – thus eluding the ephemeral nature of this kind of digital content and the ensuing time-related difficulty that it entails. Hence, it makes it possible to collect a high number of Stories at once and store them in a dedicated database. In addition, from an ethical perspective, YouTube presents far less issues of access, since it does not require to bypass the restrictions that pertain to the Instagram API. The YouTube platform is arguably more open to forms of digital methods, and its practices of use in digital social research are well-established and commonly accepted (Rieder et al., 2018). Third, YouTube provides researchers with metadata that are not otherwise provided by the Instagram platform – in particular, the reactions (e.g. likes, comments) to each (replicated) Story. Yet, as we are about to see, YouTube also has the evident disadvantage of gathering

Video Category Label	Video title occurrences, 'stories'
People & Blogs	871
How to & Style	266
Education	150
Science & Technology	149
Entertainment	142
Film & Animation	79
Music	22
Travel & Events	14
Gaming	14
Comedy	12
News & Politics	7
Sports	2
Pets & Animals	2
Autos & Vehicles	2
Total	**1732**

Figure 4. Video Title occurrences, 'stories', per Category.

Stories that, for the most part, were not posted by ordinary users. On the contrary, most of the content tagged as 'Instagram Stories' on YouTube consists of videos featuring celebrities and/or influencers, or tutorials on how to professionally produce Instagram Stories. Nonetheless, although not posted by ordinary users, YouTube allows to understand which kind of Stories ordinary users deemed relevant and worthy of being archived, and to follow them as digital objects in a more aggregated manner, as opposed to Instagram which at present allows for a more user-centred approach.

Data collection and analysis

Using the Video List module that is part of the YouTube Data Tools (Rieder, 2015), we performed a data collection starting from the keyword 'Instagram Stories', deployed in May 2019 (5 iterations). This enabled us to obtain a dataset of 11,669 videos related to this initial query. At first inspection, these include a variety of materials, not only Instagram Stories, such as music videoclips, excerpts of TV programmes, and much more. This is indicative of how the 'Instagram Stories' tag is used as a popular marketing hook to make sure content is found on YouTube.

To browse this variegated dataset, we extrapolated only those items where the occurrence 'stories' was present in the video and/or channel title. This rendered a dataset of 1732 videos where the occurrence 'stories' appears in the video title, and 562 videos where this entry appears in the channel title. Only 365 videos appear in both datasets.

Looking at video titles, these reveal a prominence of content from the People and Blogs category (Figure 4), mainly consisting in how-to content and marketing-related videos (Figure 5). Looking instead at channels, we can observe a prominence of replicated 'Story' content by celebrities and/or influencers (Figure 6), in particular, in our case, of pop icon Chiara Ferragni (Figure 7).

Alongside performing qualitative content analysis (Altheide, 1987) in order to inductively analyse the content of these videos, the use of YouTube provides the unique advantage of accessing metadata that is not otherwise retrievable via the Instagram platform

Title	Channel Title	Category	View Count	Like Count	Dislike Count	Comment Count
NOS MEILLEURES STORIES (Best Of Insta/Snap)	Mcfly & Carlito	Entertainment	4299931	166316	3603	4831
10 Instagram Stories TIPS TRICKS & HACKS | That ACTUALLY Work	Hayls World	How to & Style	3778320	70160	3285	1647
Stories	Unbox Therapy	Science & Technology	3415576	69259	2268	8841
STORIES	Porta dos Fundos	Entertainment	3381125	118217	6935	1484
FilterCopy | Honest Instagram Stories | Ft. Kritika Avasthi and Rohan Shah	FilterCopy	Comedy	2876869	71771	7734	2160
What can you do with Instagram Stories? | Instagram Business	Facebook Business	Science & Technology	2658287	19	9	n/a
TRUCOS para INSTA STORIES más COOL - Instahacks | Kika Nieto	Kika Nieto	How to & Style	2648603	149286	3982	9471
14 TRUCOS para INSTA STORIES que NO SABÍAS | @paulagonu	Paula Gonu	People & Blogs	2151049	167213	3333	6034
Respondiendo preguntas de INSTAGRAM STORIES I Kika Nieto	Kika Nieto	How to & Style	2110606	136162	8531	20957
10 Tips Instagram Stories Kekinian #KOMPAScom	TEKNO KOMPAScom	How to & Style	2092729	23483	1028	379

Figure 5. Video Title occurrences, 'stories', ordered per view count, top 10.

but may be equally (if not more) illustrative of the meaning, uses and circulation of Instagram Stories. This is particularly interesting insofar as reactions to Instagram Stories are not publicly visible on Instagram, but are only accessible to the creator of the Story.

For instance, we can analyse the comments under the replicated Instagram Stories, and thus observe the reactions of users to the content they watch. As an example, again by using the YouTube Data Tools we retrieved the comments pertaining to the most viewed video in our sample of replicated Stories, entitled 'La nascita di Leo' (English translation: 'The birth of Leo'). This is a collection of Stories by Fedez (the husband of pop icon Chiara Ferragni) that depict the birth of their son Leo. The video has 901 comments, the vast majority in Italian language. These can be analysed in various ways. In this example, we first performed an exploratory qualitative analysis of the comments ordered on the basis of the 'reply count' feature. This allows us to highlight 'controversial' comments that spark conversation among users, and thus rapidly take a glance at the nature of the ensuing debate. The analysis suggests the presence of a highly polarised discussion around the practice of documenting everyday life by celebrities using Instagram Stories. Comments with more replies seem to largely split between supportive (users congratulating the couple) and disapproving ones (users criticising the exposure of

Channel Title	Number of videos from each Channel included in the sample
Instagram Stories Replay	239
IG STORIES The Ferragnez	87
Stories	52
Element stories	49
Instagram stories	29
Fedez Stories	25
MontanaBlack Stories	22
The best of IG stories	18
Instagram Stories Live	8
Youtubers Instastories	7

Figure 6. Channel Title occurrences, 'stories', top 10.

Label	Channel Title	Category	View Count	Like Count	Comment Count
La nascita di Leo	Fedez Stories	People & Blogs	1807159	32956	901
DHASIA WEZKA EMBARAZADA? PROBLEMAS AL COMPRAR ROPA Y NOCHE ROMÁNTICA	Youtubers Instastories	People & Blogs	1152979	10884	403
Andando a Sardegna.. 😍	Fedez Stories	People & Blogs	707000	11252	880
YOSSTOP LE QUITAN LAS MUELAS DEL JUICIO 😂 Y PEL3A CON SU LORO	Youtubers Instastories	People & Blogs	701414	10776	422
DHASIA WEZKA LLORA AL ENCONTRAR A UN PERRO ABANDONADO Y LO RESCATA	Youtubers I nstastories	People & Blogs	616106	19986	623
EL MEJOR TUTORIAL NUNCA ANTES VISTO DE DHASIA WEZKA	CHISMESTORIES NETWORK	People & Blogs	603362	13375	223
CHIARA FERRAGNI E FEDEZ TOUR COMPLETO DELLA NUOVA CASA VIDEO UFFICIALE	Stories	People & Blogs	580702	5583	593
20.000€ von der Bank abholen😱 MontanaBlack Instagram Story	MontanaBlack Stories	People & Blogs	560362	7226	869
Lambo abgeholt😍 erste Polizeikontrolle😱 MontanaBlack Instagram Story	MontanaBlack Stories	People & Blogs	547890	9416	1074
GRAZIE DI TUTTO. ANITA STORIES	Anita Stories	How to & Style	509962	29117	11079

Figure 7. Video list from Channel Title occurrences, 'stories', ordered per view count, top 10.

Topic	1	2	3	Total
Celebrity	35,03%	33,65%	5,15%	25,34%
Documentation	6,35%	45,97%	3,68%	15,00%
Event	14,72%	8,06%	1,47%	8,98%
Baby	35,79%	2,84%	5,51%	18,41%
Other	2,79%	3,32%	80,51%	27,27%
Parenthood	5,33%	6,16%	3,68%	5,00%
Total (general)	**100,00%**	**100,00%**	**100,00%**	**100,00%**

Figure 8. Topic analysis, comments to 'La nascita di Leo' (percentage values, filtered per sentiment).

the child). To corroborate this insight, we performed a sentiment analysis of all the comments, which have been manually coded using: 1 = positive; 2 = negative; 3 = neutral/unrelated. Overall, the analysis confirms this polarisation. The relative majority of the sample is made of positive reactions (44.93%); yet, there is also a considerable presence of negative comments (24.06%).

Alongside the sentiment analysis, we also performed a topic analysis by manually coding each of these comments on the basis of the topic it deals with. If we filter the topic analysis per sentiment (Figure 8), we can evidence the narrations that accompany this conversation, which split between positive comments about the couple (here tagged as 'celebrity' content) and the baby (tagged as 'baby'), and negative comments about the use of Instagram Stories to document the child's birth (here tagged as 'documentation'). These criticise the way the couple rendered this event a piece of showbiz, as in the example below:

> *yeah but it is not normal that you are giving birth and you only think about filming it* [5]

Others, instead, defend the choice of the couple to document the birth of their child on Instagram Stories:

> *I find this anger totally out of place. 'They are the only ones who had a child . . .' If you are annoyed by the fact they publicly disclose something, you should not be here commenting and giving views, it is called coherence. I never liked Fedez (Ferragni's husband, ndr) as an artist but I must say that in these pictures I have seen a touching dad, and believe me he is the first one to laugh reading the silly things you write.*[6]

This kind of analysis, we argue, is quite insightful for two main reasons. On the one hand, we can illustrate how Instagram Stories are clearly understood by users as culturally contentious objects and represent a key marketing tool beyond the boundaries of the Instagram platform. On the other hand, we are able to observe and decipher ephemeral digital content as a cultural object in a way that the Instagram platform does not allow (irrespective of its API closure), enabling us to expand our understanding of Instagram Stories in their social and cultural complexity.

Final discussion, limitations and further research

The article has experimented two research strategies to 'repurpose' digital methods for the study of ephemeral digital content such as Instagram Stories, which represents an emergent, relevant and understudied dimension of present-day digital cultures. In Strategy 1, we employed scraping techniques as a means to access and observe ephemeral content in the timeframe it is made available by users. Scraping techniques enable the researcher to organise ephemeral content for cultural analysis and delve deep in its multimodality, thus offering a rich picture of this type of content as a social and cultural entity. Yet, the use of scraping techniques does not allow to exhaustively address ephemeral content as a flow, as the 'story' is extricated out of its context – an aspect which, depending on the research question, may be necessary to consider in the research design. In Strategy 2, we 'circumvented the object of study', looking at another platform – YouTube – where we find replicated ephemeral content. This allowed us to observe how digital spaces remain prominent milieus for archival practices (Burgess and Green, 2018) also of content that is designed to disappear. The study of Instagram Stories on YouTube enables researchers to investigate more closely the broader context within which ephemeral content appears, its function and perception as offered by the comments, which represent extremely valuable pieces of information otherwise unavailable on the Instagram platform.

Despite their experimentation in isolation, we contend it is the combination of these strategies that offers a rich set of insights. First, by using digital methods instead of traditional observational techniques we were able to capture and store Instagram Stories in dedicated databases, which allowed us to perform more refined qualitative analyses. Moreover, the combination of the two strategies granted us the possibility to investigate and cast a light on the underlying mechanisms of influence between platforms and users, which would have been impossible to unearth otherwise. Specifically, we saw that, although Instagram Stories encourage ordinary users to express their own creativity – a call that users genuinely embrace – they tend to do so by adhering to the Instagram platform *aesthetics* as well as to *norms* of influencer marketing. Finally, our study shows how ephemeral and archival cultures coexist across social media platforms in a seamless dimension. The study of ephemeral and archival cultures as separate and neatly distinguished entities does not allow to grasp the extent to which ephemeral and traditional content actually interact and complement each other. Our analysis of Instagram Stories, albeit primarily pointed at methodological reflections, ultimately shows that ephemeral content adds a new layer of complexity to digital cultures and their practices, insofar as it does not suppress the cultural logics of archival but engages with it, creating a mix of new and established practices and forms of sociality.

Furthermore, it is worth acknowledging that, however useful, our contribution is not without limitations, which concern both our global methodological framework and the two specific strategies. Regarding the general framework, we acknowledge the limited number of cases our empirical research draws on. This was due to the exploratory nature of the project. Also, we have not paid specific attention to the practice of re-sharing content via Instagram Stories, which is increasingly popular in a number of social and cultural contexts. While this was not made object of analysis in our article, it certainly represents a relevant phenomenon, which should be investigated further in future

research. Furthermore, we focus only on two social media platforms; it is our conviction that a cross-platform approach (Rogers, 2019), for instance considering also Twitter, Pinterest or TikTok, could deepen our understanding of Instagram Stories as technical and cultural devices as well as of the tension between ephemeral and archive cultures. More specifically, concerning Strategy 1, for how systematic and transparent we acknowledge that our procedure of identifying ordinary users and content on Instagram might sound somewhat arbitrary. As our research project is the first of this kind, we needed an initial point of entry into the Instagram Stories ecosystem. Yet, further big data and statistical analysis could help in devising more objective data collection processes, for example, by (a) identifying ad hoc hashtags and/or public Instagram pages that primarily aggregate ordinary Stories; (b) automatising the procedure to detect bots, fan pages and influencers; (c) estimating the average number of followers of ordinary users. Strategy 2, in turn, presents two main limits. First, collecting Stories through YouTube means primarily running into celebrity or influencers' Stories – although not directly posted by these actors. Second, while YouTube allows researchers to access the public debate around Instagram Stories, it is not possible to be sure that those reacting to and commenting Stories-related videos are also users that actually and routinely consume Instagram Stories or have a particular interest in them. Again, big data and statistical analysis can be helpful here. For example, a large-scale mapping of YouTube – perhaps also using the YouTube Data Tools – can help researchers to detect specific communities of Instagram users that are posting and discussing ordinary Stories.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Lucia Bainotti https://orcid.org/0000-0003-4613-8841

Alessandro Caliandro https://orcid.org/0000-0002-1168-882X

Alessandro Gandini https://orcid.org/0000-0002-7705-7625

Notes

1. See https://www.instagram.com/developer/changelog/ (Last accessed 16 November 2019).
2. See https://www.storysaver.net/ (Last accessed 10 December 2019).
3. See https://www.instagram.com/about/legal/terms/api/ (Last accessed 22 May 2020).
4. See https://help.instagram.com/151273688993748 (Last accessed 10 December 2019).
5. Original comment in Italian (translated by the authors).
6. Original comment in Italian (translated by the authors).

References

Altheide DL (1987) Reflections: ethnographic content analysis. *Qualitative Sociology* 10(1): 65–77.

Banks M (2007) *Using Visual Data in Qualitative Research*. London: SAGE.

Bayer JB, Ellison NB, Schoenebeck SY, et al. (2016) Sharing the small moments: ephemeral social interaction on Snapchat. *Information, Communication & Society* 19(7): 956–977.

Böhme R and Köpsell S (2010) Trained to accept? A field experiment on consent dialogs. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, Atlanta, GA, 10–15 April, pp. 2403–2406. New York: ACM.

Bruns A (2018) Facebook shuts the gate after the horse has bolted, and hurts real research in the process. *Internet Policy Review*. Available at: https://policyreview.info/articles/news/facebookshuts-gate-after-horse-has-bolted-and-hurts-real-research-process/786 (accessed 8 February 2020).

Bruns A (2019) After the 'APIcalypse': social media platforms and their fight against critical scholarly research. *Information, Communication & Society* 22(11): 1544–1566.

Bucher T (2017) The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society* 20(1): 30–44.

Bucher T (2018) *If . . . then: Algorithmic Power and Politics*. Oxford: Oxford University Press.

Burgess J and Green J (2018) *Youtube: Online Video and Participatory Culture*. Hoboken, NJ: John Wiley & Sons.

Caliandro A and Gandini A (2017) *Qualitative Research in Digital Environments: A Research Toolkit*. London: Routledge.

Carah N and Shaul M (2016) Brands and Instagram: point, tap, swipe, glance. *Mobile Media & Communication* 4(1): 69–84.

Chellapilla K, Larson K, Simard PY, et al. (2005) Computers beat humans at single character recognition in reading based Human Interaction Proofs (HIPs). *CEAS*. Available at: https://www.semanticscholar.org/paper/Computers-beat-Humans-at-Single-Character-in-based-Chellapilla-Larson/e405e0f1bca55b7d0e17656533f63fd025c2dd58 (accessed 29 April 2020).

Constine J (2018) Stories are about to surpass feed sharing. Now what? *Techcrunch*. Available at: https://techcrunch.com/2018/05/02/stories-are-about-to-surpass-feed-sharing-now-what/ (accessed 10 February 2020).

Cooley R, Mobasher B and Srivastava J (1997) Web mining: information and pattern discovery on the World Wide Web. In: *Proceedings of the ninth IEEE international conference on tools with artificial intelligence*, Newport Beach, CA, 3–8 November, pp. 558–567. New York: IEEE.

Fiesler C, Beard N and Keegan BC (2020) No robots, spiders, or scrapers: legal and ethical regulation of data collection methods in social media terms of service. *cmci.colorado.edu*. Available at: https://cmci.colorado.edu/~cafi5706/ICWSM2020_datascraping.pdf (accessed 29 April 2020).

Fiesler C, Lampe C and Bruckman AS (2016) Reality and perception of copyright terms of service for online content creation. In: *CSCW'16: Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*, San Francisco, CA, 27 February–2 March, pp. 1450–1461. New York: ACM.

Franzke AS, Bechmann A, Zimmer M, et al. (2019) Internet research: ethical guidelines 3.0: association of internet researchers. Available at: https://aoir.org/reports/ethics3.pdf

Geboers M (2019) 'Writing' oneself into tragedy: visual user practices and spectatorship of the Alan Kurdi images on Instagram. *Visual Communication*. Epub ahead of print 26 June. DOI: 10.1177/1470357219857118.

Georgakopoulou A (2017) 17 small stories research: a narrative paradigm for the analysis of social media. In: Sloan L and Quan-Haase A (eds) *The SAGE Handbook of Social Media Research Methods*. London: SAGE, pp. 266–281.

Gerlitz C and Rieder B (2018) Tweets are not created equal: investigating Twitter's client ecosystem. *International Journal of Communication* 12: 528–547.

Gerrard Y (2018) Beyond the hashtag: circumventing content moderation on social media. *New Media & Society* 20(12): 4492–4511.

Gibbs M, Meese J, Arnold M, et al. (2015) # Funeral and Instagram: death, social media, and platform vernacular. *Information, Communication & Society* 18(3): 255–268.

Hagen S (2018) Rendering legible the ephemerality of 4chan/pol/. *Oilab*. Available at: https://oilab.eu/rendering-legible-the-ephemerality-of-4chanpol/ (accessed 10 February 2020).

Handyside S and Ringrose J (2017) Snapchat memory and youth digital sexual cultures: mediated temporality, duration and affect. *Journal of Gender Studies* 26(3): 347–360.

Heath C, Hindmarsh J and Luff P (2010) *Video in Qualitative Research*. London: SAGE.

Hess A (2015) The selfie assemblage. *International Journal of Communication* 9: 1629–1646.

Hu Y, Manikonda L and Kambhampati S (2014) What we Instagram: a first analysis of instagram photo content and user types. In: *Proceedings of the eighth international AAAI conference on weblogs and social media*, Ann Arbor, MI, 1–4 June.

Instagram (2016) Introducing Instagram stories. Available at: http://instagram.tumblr.com/post/148348940287/160802-stories (accessed 10 February 2020).

Jurgenson N (2013) Pics and it didn't happen. *The New Inquiry*. Available at: http://thenewinquiry.com/essays/pics-and-it-didnt-happen/ (accessed 10 February 2020).

Kazansky B, Torres G, van der Velden L, et al. (2019) Data for the social good: toward a data-activist research agenda. In: Daly A, Devitt SK and Mann M (eds) *Good Data*. Amsterdam: Institute of Network Cultures, pp. 244–259.

Koefed J and Larson MC (2016) A snap of intimacy: photo-sharing practices among young people on social media. *First Monday* 21(11): 1–9.

Landers RN, Brusso RC, Cavanaugh KJ, et al. (2016) A primer on theory-driven web scraping: automatic extraction of big data from the Internet for use in psychological research. *Psychological Methods* 21(4): 475–492.

Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Leaver T, Highfield T and Abidin C (2020) *Instagram: Visual Social Media Cultures*. Cambridge; Medford, MA: Polity Press.

Light B, Burgess J and Duguay S (2018) The walkthrough method: an approach to the study of apps. *New Media & Society* 20(3): 881–900.

Manovich L (2016) Instagram and contemporary image. *Manovich.net*. Available at: http://manovich.net/index.php/projects/instagram-and-contemporary-image (accessed 8 February 2020).

Marwick AE (2015) Instafame: luxury selfies in the attention economy. *Public Culture* 27(1(75): 137–160.

Mayer-Schönberger V (2009) *Delete: The Virtue of Forgetting in the Digital Age*. Princeton, NJ: Princeton University Press.

Nashmi EA and Painter DL (2018) Oh snap: chat style in the 2016 US presidential primaries. *Journal of Creative Communications* 13(1): 17–33.

Newberry C (2019) 37 Instagram Stats that matter to marketers in 2020. *Hootsuite*. Available at: https://blog.hootsuite.com/instagram-statistics/ (accessed 10 February 2020).

Page R (2013) *Stories and Social Media: Identities and Interaction*. London: Routledge.

Page R (2015) The narrative dimensions of social media storytelling. In: De Fina A and Georgakopoulou A (eds) *The Handbook of Narrative Analysis*. Hoboken, NJ: John Wiley & Sons, pp. 329–347.

Perriam J, Birkbak A and Freeman A (2019) Digital methods in a post-API environment. *International Journal of Social Research Methodology*. Epub ahead of print 25 October. DOI: 10.1080/13645579.2019.1682840.

Postill J and Pink S (2012) Social media ethnography: the digital researcher in a messy Web. *Media International Australia* 145: 123–134.

Puschmann C (2019) An end to the wild west of social media research: a response to Axel Bruns. *Information, Communication & Society* 22(11): 1582–1589.

Reidenberg JR, Breaux T, Cranor LF, et al. (2015) Disagreeable privacy policies: mismatches between meaning and users' understanding. *Berkeley Technology Law Journal* 30(1): 39–68.
Rieder B (2015) Introducing the YouTube data tools. *The Politics of Systems*. Available at: http://thepoliticsofsystems.net/2015/05/exploring-youtube/ (accessed 8 February 2020).
Rieder B, Matamoros-Fernández A and Coromina Ò (2018) From ranking algorithms to 'ranking cultures' Investigating the modulation of visibility in YouTube search results. *Convergence* 24(1): 50–68.
Roesner F, Gill BT and Kohno T (2014) Sex, lies, or kittens? Investigating the use of Snapchat's self-destructing messages. In: *Proceedings of the financial cryptography and data security conference*, Christ Church, 3–7 March.
Rogers R (2013) *Digital Methods*. Cambridge, MA: MIT Press.
Rogers R (2019) *Doing Digital Methods*. London: SAGE.
Rokka J and Canniford R (2016) Heterotopian selfies: how social media destabilises brand assemblages. *European Journal of Marketing* 50(9–10): 1789–1813.
Rose G (2016) *Visual Methodologies: An Introduction to Researching with Visual Materials*. London: SAGE.
Senft TM and Baym NK (2015) What does the selfie say? Investigating a global phenomenon. *International Journal of Communication* 9: 1588–1606.
Vázquez-Herrero J, Direito-Rebollal S and López-García X (2019) Ephemeral journalism: news distribution through Instagram Stories. *Social Media + Society* 5(4): 1–13.
Venturini T and Rogers R (2019) 'API-based research' or how can digital sociology and journalism studies learn from the Facebook and Cambridge analytica data breach. *Digital Journalism* 7(4): 532–540.
Von Ahn L, Blum M, Hopper NJ, et al. (2003) CAPTCHA: using hard AI problems for security. *EUROCRYPT* 2003: 297–311.
Warren J (2019) 10 strategies to drive traffic and sales from Instagram. *Later*. Available at: https://later.com/blog/drive-traffic-from-instagram/ (accessed 8 February 2020).
Waterman T (2020) Web scraping is now legal. *Towards Data Science*. Available at: https://towardsdatascience.com/web-scraping-is-now-legal-6bf0e5730a78 (accessed 8 February 2020).
Weltevrede E (2016) *Repurposing digital methods: the research affordances of platforms and engines*. PhD Dissertation, University of Amsterdam, Amsterdam.
Zimmer M (2010) 'But the data is already public': on the ethics of research in Facebook. *Ethics and Information Technology* 12(4): 313–325.
Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.

Author biographies

Lucia Bainotti is a PhD student at the Network for the Advancement of Social and Political Studies, University of Turin and University of Milan. Her research interests include digital methods, digital consumer culture, self-branding and gender.

Alessandro Caliandro (PhD, University of Milan) is a senior lecturer in the Department of Political and Social Sciences at University of Pavia. His current research focuses on digital methods, digital consumer culture, social media affordances, and smartphone use.

Alessandro Gandini is senior lecturer in cultural sociology at the University of Milan, Department of Social and Political Sciences. His research interests concern the transformations of work in the digital society, digital methods and consumer cultures, the critical study of algorithms and digital platforms.